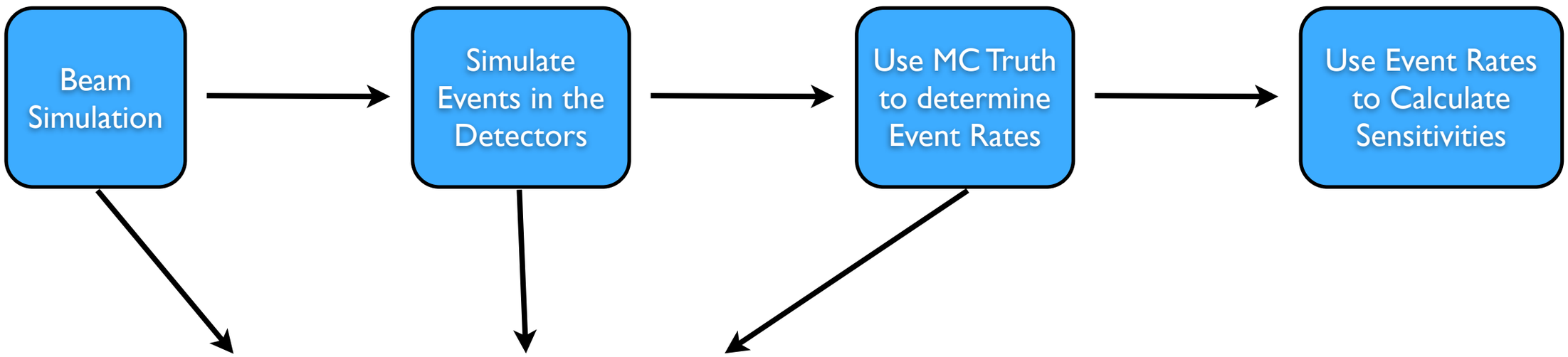
LAr1 Sensitivity Calculations

From Beam to Plots

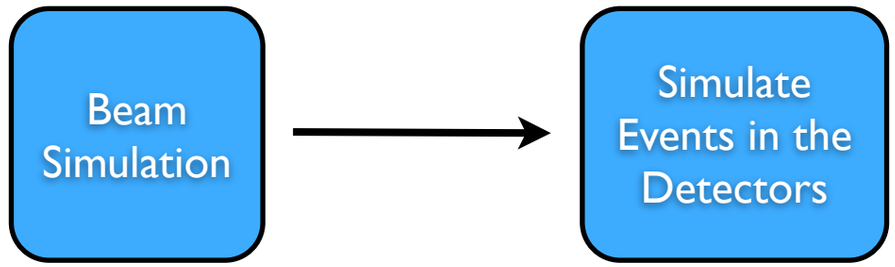
Corey Adams

Overview of the path to sensitivities:



The entire chain of events is performed by several people (Zarko, Georgia, Roxanne, Corey, and sometimes more) and so the code is scattered throughout the /uboone directory. I've tried to collect the locations as much as I can so that a newcomer could pick up what we've done without having to search too much.

Please stop me for questions, comments, or corrections. I'm going to post an updated version of this after the meeting with all of the comments included.



Start with MiniBoone prediction of Particles at the source.

Using gsimple, we take the particles and propagate them (and in particular the neutrinos) to a detector location. We then freeze the simulation here and create flux files representing the neutrino flux at each detector.

Using the flux files (we create 1000 or 2000 per baseline and beam mode), we use GENIE to simulate the interactions inside the detector. This too gets frozen and written out.

We feed the GENIE output into larg4 to propagate the particles through the detectors. This is useful because it helps us determine rates of various backgrounds.

Again, we write out the particles and state of the simulation after larg4 to be used in the next step.

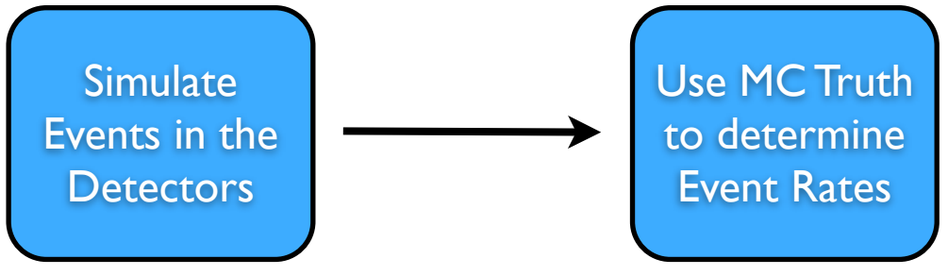
Flux Files at: /uboone/data/bnb_gsimple_fluxes_*
Genie Generation Code: /uboone/app/users/guenette/Gen_uboone/condor/
Also, code in .../guenette/Gen_LAr1/condor/
Grid scripts in the same guenette folder
larg4 code: /uboone/app/users/cadams/frozen/ART_fhicl/larg4_and_daq.fcl
Grid scrips in same cadams folder

Important Notes for the previous steps:

We create 4 sets of genie files for each detector. We have neutrino and antineutrino mode, as well as the fully oscillated versions of neutrino and antineutrino mode. The full osc events allow us to simulate signal.

The only input files the previous steps require (other than the output from the previous step in the chain) are the correct geometry files. Everything else is standard larsoft*. All the geometry files can be found at:
/uboone/app/users/cadams/frozen/Geometry/gdml (among other places)

*The exception to the rule: we use a slightly modified Geometry package that does not print "Can't find nearest channel" exceptions. This is thrown when the daq simulator can't locate a wire to deposit charge on. Normally this isn't a problem, but our LAr1 geometry files don't have wires and this exception gets thrown a LOT. It generates 7 gigabytes of logfile per input file, and after 2000 input files this is 14 terabytes (!!!!) of logfiles. Considering that /uboone/data is 40 TB in capacity, this can be bad.



Have events in each detector, use a LArSoft Analyzer module to look at each event and determine:

NC/CC, nu/nubar, energy, neutrino species, travel length, decay parent, interaction vertex, parent momentum, parent vertex, angles of outgoing lepton, and location information about produced gammas if there are any. We can simulate reconstruction of neutrino energy through smearing, but we haven't been using this yet.

We make a fiducial cut on the CC events, but are currently keeping all NC Events.

Each Ana module produces a root file with a tree (ntuple) of this information for each event. We run this over all the events, and combine all of the ntuples.

At this stage, we also extract the POT per file (since each file is generated with 1000 or 2000 events, POT varies at each file).

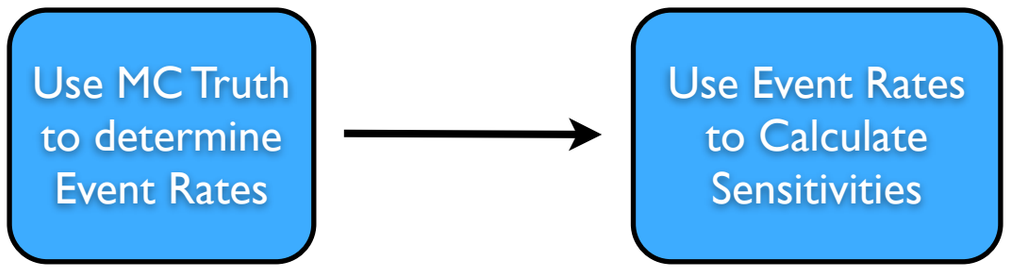
All of the ntuples are combined via a root macro

Code: /uboone/app/users/cadams/frozen/GENIEextractor/NuAna.h, NuAna.cxx
Combination marco: /uboone/app/users/cadams/frozen/TTrees_to_TChain.C

Event Rate Details

We have a comparison of event rates here: https://docs.google.com/spreadsheet/ccc?key=0AuustSx14WrmdC1xcWxCWmxEMzZzMXFpTXBiQXQtblE#gid=0

Our ratio of events/POT/Ton for nu to nubar is about 4 to 1. We think that this is reasonable, but the nubar rate is much worse than previous results. This affects the current sensitivities in nubar mode badly, and I think it is a big reason we haven't reproduced the results shown in the white paper.



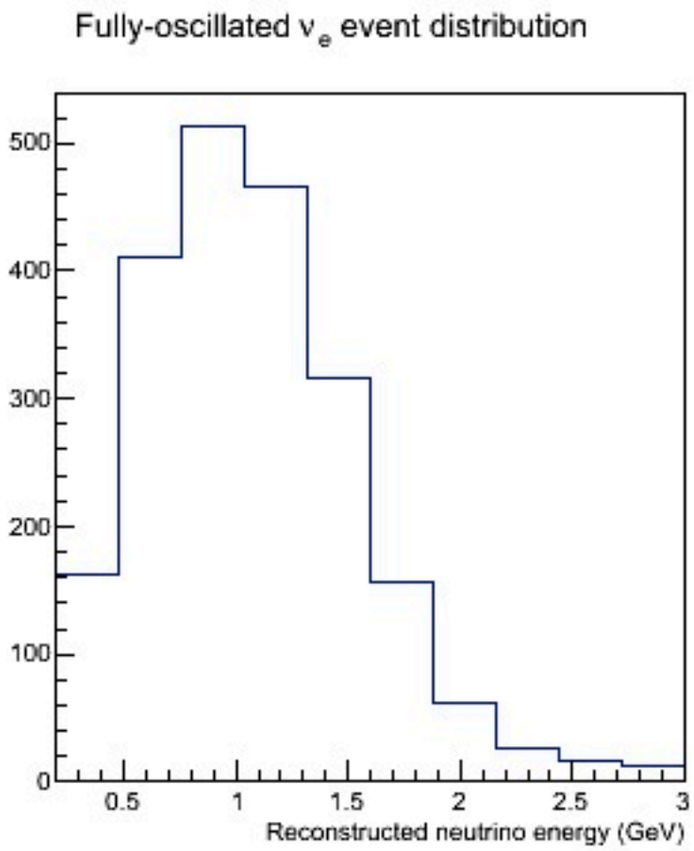
With ntuples in hand (usually ~ 10^6 events per file), we can count the number of events of each type (meaning, CC or NC, nue, numu, etc).

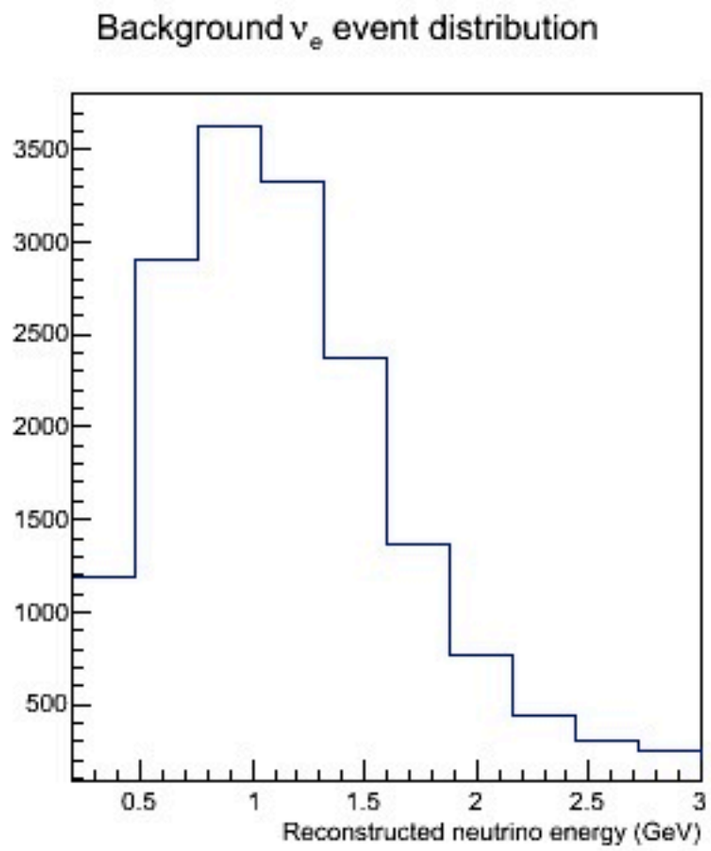
The events rates may need to be corrected to account for things like flux reweighing. The same code can be used to apply reconstruction efficiencies or misID rates to events. (Comment: we still need to decide on the misID rates, and they have a big effect on background when looking for nue appearance.)

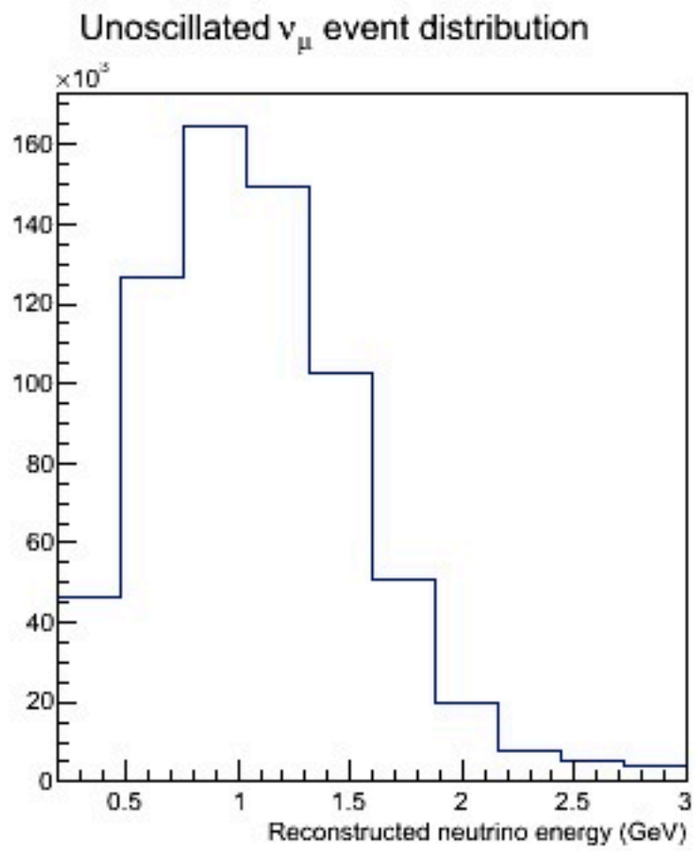
In practice, the way we do this is to give each event a weight and efficiency. Then, in summing event rates, we add something like (1 event from simulation)*(weight correction)* (efficiency of reconstruction/misID)

There is a root macro using Georgia's implementation of Combined Fit to use the event rates to determine an experimental sensitivity. More on this from Georgia, hopefully.

Flux code: /uboone/app/users/cadams/frozen/CombinedFitting/reprocessing/ntuple_*.C and .h
This code can (and is) frequently modified to use varying efficiencies, but the flux reweighting is constant.







This is what the event rate histograms look like, and what we use for the combined fitting method. (Neutrino mode at 700m)

Technical Notes on Running on the Grid

In a folder like /uboone/app/users/cadams/frozen/ART_fhicl, there are several .fcl files (such as larg4_and_daq.fcl) that are very generic to run larsoft for this analysis. There are many *.sh files that are grid submission scripts that fill in the pieces to these .fcl files and submit them to the grid.

For someone else to come along and use these grid scripts, I'd recommend you copy the whole folder to your home directory and then you can edit HOW_TO_RUN line 17 to use YOUR grid proxy instead of mine.

There is a section in the grid scripts for modifications that's clearly bookended by "###---------- ....", the changes to switch from one configuration to another *should* all happen there.

Launch the jobs with:
source HOW_TO_RUN yourcondorscript.sh njobs /path/to/output/directory

It's important to have "umask 002" at the start of the grid script to allow us to modify the written files, since they are written by user "ubooneana"

Use Event Rates to Calculate Sensitivities

There are 3 versions of the code to do this job, for 3 distinct reasons.

One version, georgias_multiple_detector_fit.C, will take the raw ntuples without any efficiency or reweighting and use them to calculate sensitivities.

The second version, multiple_detector_fit.C, will use ntuples that are reweighted and have efficiencies or misIDs applied to the events.

The last version, covariance_matrix_fit.C, uses the same ntuples as the previous version but has the machinery built in to read a covariance matrix (or several) as the systematic error.

All of these programs produce (and save) the sensitivity plots they generate. Further, reading the ntuples and creating histograms of event rates per energy bin is a slow process. So, these programs are written to save this information, but automatically rerun if the ntuples are newer than the event rate files.

Code for Sensitivity Calculations, in /uboone/app/users/cadams/frozen/CombinedFitting/
georgias_multiple_detector_fit.C
multiple_detector_fit.C
covariance_matrix_fit.C

Generating Plots

Each version of the fitting code has a few parameters to adjust for different plots. You can adjust the number of sampling points on the sensitivity plots to adjust the resolution (more points = better resolution, but runtime scales as n^2).

You can change the variable to point to a different folder of ntuples to run with different efficiencies applied, and you can use a scaling factor to play with the fiducial volume of the detectors (this is a crude but very quick way to see the effect of changing the detector size).

To fiddle with the plot settings themselves, jump straight to the bottom of the code.
There is also another macro that reads the saved graphs from the fitting routine and can be used to put the sensitivity curves on the same graph for direct and easy comparison.

Macro for plotting sensitivities on the same plot: /uboone/data/users/cadams/frozen/CombinedFitting/plotMaker.C